Solvent molecules bridge the mechanical unfolding transition state of a protein

Lorna Dougan*†, Gang Feng‡, Hui Lu‡, and Julio M. Fernandez*†

*Department of Biological Sciences, Columbia University, New York, NY 10027; and ‡Department of Bioengineering, University of Illinois, Chicago, IL 60607

We demonstrate a combination of single molecule force spectroscopy and solvent substitution that captures the presence of solvent molecules in the transition state structure. We measure the effect of solvent substitution on the rate of unfolding of the I27 titin module, placed under a constant stretching force. From the force dependency of the unfolding rate, we determine $\Delta x_{\text{u}}$, the distance to the transition state. Unfolding the I27 protein in water gives a $\Delta x_{\text{u}} = 2.5$ Å, a distance that compares well to the size of a water molecule. Although the height of the activation energy barrier to unfolding is greatly increased in both glycerol and deuterium oxide solutions, $\Delta x_{\text{u}}$ depends on the size of the solvent molecules. Upon replacement of water by increasing amounts of the larger glycerol molecules, $\Delta x_{\text{u}}$ increases rapidly and plateaus at its maximum value of 4.4 Å. In contrast, replacement of water by the similarly sized deuterium oxide does not change the value of $\Delta x_{\text{u}}$. From these results we estimate that six to eight water molecules form part of the unfolding transition state structure of the I27 protein, and that the presence of just one glycerol molecule in the transition state is enough to lengthen $\Delta x_{\text{u}}$. Our results show that solvent composition is important for the mechanical function of proteins. Furthermore, given that solvent composition is actively regulated in vivo, it may represent an important modulatory pathway for the regulation of tissue elasticity and other important functions in cellular mechanics.

**Water** is recognized as an active participant in biological processes, responding structurally and dynamically to the presence of other molecules in subtle and nonintuitive ways (1, 2). In particular, water is thought to play a dynamic role in protein folding and unfolding (3, 4). Water facilitates the necessary changes of the hydrogen bonding network, allowing fast conformational changes (5, 6). During protein folding, water mediates the collapse of the chain and search for native state conformations, contributing to both enthalpic and entropic stabilization (4). An understanding of the mechanisms of protein unfolding and refolding must incorporate the solvating environment which envelopes the molecule. Indeed, the energy landscape of a protein can be affected by changing the solvent environment properties, suggesting different structures and altering the folding thermodynamics and kinetics (7, 8). There have been a number of efforts, using bulk experimental techniques, to understand the influence of the solvent environment on the behavior of proteins (7–15). Although such experiments have revealed a wealth of information regarding the thermodynamics of protein folding, very little is known about the role that solvent molecules play on the structure of the folding/unfolding transition state of a protein, which is the main determinant of protein dynamics.

Single-molecule force spectroscopy is an excellent tool to probe transition states in a protein (16–22). This technique is used to apply a mechanical force to a single protein, causing the protein to unfold and extend along a well defined reaction coordinate; the end-to-end length of the protein. Along this unfolding pathway, a mechanically resistant transition state determines the force-dependent rate of unfolding, $k_u(F)$, easily measured with force spectroscopy techniques (16, 19, 21). The force dependency of the unfolding rate is typically fit with a straightforward Arrhenius term that measures properties of the unfolding transition state (19). In its simplest representation, the unfolding transition state is characterized by two parameters: the size of its activation energy, $\Delta G_{\text{u}}$, and the elongation of the protein necessary to reach the transition state, $\Delta x_{\text{u}}$. Of particular interest are the force spectroscopy measurements of $\Delta x$ which provide a direct measure of the length scales of a transition state, which were hitherto unknown. For example, for protein folding the distance to the folding transition state, $\Delta x_{\text{f}}$, was found to be between 8 Å (17) and 60 Å (23), in rough agreement with the expected role of long range hydrophobic forces (24). For protein unfolding $\Delta x_{\text{u}}$ was found to be much shorter, in the range of 1.7–2.5 Å (16, 19, 21). These values of $\Delta x_{\text{u}}$ are comparable to the size of a water molecule (25), suggesting that water molecules are integral components of the unfolding transition state.

Here we use single molecule force-clamp spectroscopy to test this prediction by measuring the distance to the unfolding transition state, $\Delta x_{\text{u}}$, of the human cardiac titin domain I27 in the presence of glycerol and deuterium oxide. Glycerol is a good hydrogen bonding molecule which is ubiquitous in living systems (26) is known to improve protein stability (9), and is larger in size than water. Deuterium oxide forms stronger hydrogen bonds than water while having a similar size (45). Our experiments directly demonstrate that solvent molecules form part of the structure of the mechanical transition state of a protein.

**Results and Discussion**

In our experiments, we construct polyproteins with eight repeats of the human cardiac titin domain I27 (27). Polyproteins are multidomain proteins composed of identical repeats of a single protein (28). The I27 protein is ideal for these experiments given that its mechanical properties have been well characterized both experimentally (21, 28–30) and also in silico using molecular dynamics techniques (31–33). The use of polyproteins is advantageous in that they provide a clear mechanical fingerprint to distinguish them against a background of spurious interactions and also provide us with a larger number of events per recording than otherwise possible with monomers (21).

We measure the properties of the mechanical unfolding transition state of the I27 protein by measuring the force dependency of the unfolding rate of single I27$_n$ polyproteins. When a protein is subjected to an external force its unfolding rate, $k_u$, is well described by an Arrhenius term of the form $k_u(F) = k_u^0 \exp(F\Delta x_{\text{u}}/k_BT)$ (16, 19, 34) where $k_u^0$ is the unfolding rate in the absence of external forces, $F$ is the applied force and

Author contributions: L.D., H.L., and J.M.F. designed research; L.D. and G.F. performed research; L.D. and H.L. contributed new reagents/analytic tools; L.D. and G.F. analyzed data; and L.D., H.L., and J.M.F. wrote the paper.

The authors declare no conflict of interest.

This article contains supporting information online at www.pnas.org/cgi/content/full/0706075105/DC1.

© 2008 by The National Academy of Sciences of the USA

PNAS | March 4, 2008 | vol. 105 | no. 9 | 3185–3190

www.pnas.org/cgi/doi/10.1073/pnas.0706075105
obtained unfolding traces, shown in Fig. 1, decreased the value of $k_u$ by glycerol has a large effect on the force dependency of glycerol (filled circles). These data showed that replacing water of the unfolding rate in 20% glycerol (filled triangles) and 30% glycerol (Fig. 1B Top). Furthermore, we obtain an estimate of the standard error of $k_u(F)$, using the bootstrapping technique (34, 35). We repeated these measurements over the force range between 120 and 220 pN and obtained the force-dependency of the unfolding rate in standard aqueous solution (Fig. 1C, filled squares). We fit the Arrhenius rate equation to the data (Fig. 1C, solid line over filled squares), and obtained $k_u^0 = 1.25 \times 10^{-4} \pm 10^{-5}$ s$^{-1}$ and $\Delta x_u = 2.5 \pm 0.1$ Å. From the measured value of $k_u^0$ we can readily estimate the value of $\Delta G_u = 23.11$ kcal mol$^{-1}$.

To probe the role of the solvent in setting the structure of the unfolding transition state, we studied the effect of solvent substitution on the force dependency of the unfolding rate. Fig. 1B Middle and Bottom shows averaged unfolding traces and their corresponding exponential fits obtained at different forces in solutions containing 20% glycerol (Fig. 1B Middle) and 30% glycerol (Fig. 1B Bottom). Fig. 1C shows the force dependency of the unfolding rate in 20% glycerol (filled triangles) and 30% glycerol (filled circles). These data showed that replacing water by glycerol has a large effect on the force dependency of unfolding. It is readily apparent that the introduction of glycerol decreased the value of $k_u^0$ (larger $\Delta G_u$) while increasing the slope $(\Delta x_u/k_u T)$. Fits of the Arrhenius term to these data measured the increase in the value of $\Delta G_u$ (26.16 kcal mol$^{-1}$ and 27.16 kcal mol$^{-1}$ in 20% and 30% glycerol, respectively). Our estimates of the unfolding activation energy, $\Delta G_u$, depend on the value of the preexponential factor, $A$. The value of $A$ is not known and, for proteins, ranges from $10^{13}$ s$^{-1}$ to values as low as $10^4$ s$^{-1}$ (36).

Here we use $10^5$ s$^{-1}$, which is used in bulk studies of solution protein biochemistry (38). Although the absolute values of $\Delta G_u$ depend on the value taken for $A$, the observation that $\Delta G_u$ increases with the concentration of glycerol remains unchanged. Although previous bulk studies have measured the stabilization of proteins using glycerol (9, 10, 26), our experiments provide a single-molecule-level demonstration of the stabilization of a protein by an osmolyte. More strikingly, the introduction of glycerol caused a large increase in the distance to transition state from $\Delta x_u = 2.5$ Å in aqueous and up to $4.0 \pm 0.1$ Å and $4.1 \pm 0.1$ Å in 20% and 30% glycerol, respectively (Fig. 1C). Earlier molecular dynamics simulations of forced unfolding of I27 suggested that when a stretching force is applied to the protein’s termini, resistance to unfolding originated from a set of hydrogen bonds between two parallel $\beta$-strands ($A'$ and $G$) of the protein structure (31–33). These $\beta$-strands must slide past one another for unfolding to occur. Because the hydrogen bonds are perpendicular to the axis of extension, they must rupture simultaneously to allow relative movement of the two termini. Hence, the rupture of these bonds defined the unfolding transition state for the I27 protein. A molecule like glycerol, larger than water (5.6 Å versus 2.5 Å, respectively) (37) but equally competent to form backbone hydrogen bonds would lead to a larger separation between the two key $\beta$-strands, enlarging the value of $\Delta x_u$. If this simple view were correct, the value of $\Delta x_u$ should increase to a value of 5.6 Å, when glycerol fully replaced water as the solvent.

The large stabilizing effect of glycerol (Fig. 1C) combined with the increased slope of the force-dependency of the unfolding rate made it difficult to do force-clamp experiments at concentrations >30%. Indeed, Fig. 1C (filled diamond) shows that the unfolding rate at 200 pN dropped 12-fold when the glycerol concentration
was increased from 30% to 50%. Our force clamp instrumentation works well measuring rates as slow as ~0.01 s⁻¹. Outside of this range, measurements become limited by cantilever drift in recordings that last longer than ~60 s. Hence, a very steep force dependency such as the one we encountered for the unfolding rate constant of Glycine in aqueous solution at a constant velocity of 400 nm s⁻¹ results in a force-extension curve with a sawtooth pattern having equally spaced force peaks of 28 nm. A typical force-extension curve is shown for I27 in aqueous solution (Top), 30% v/v vol glycerol (Middle), and 100% vol vol glycerol (Bottom). The distribution of unfolding forces obtained from Monte Carlo simulations are shown in red where the unfolding rate constant is 4.98 × 10⁻⁹ s⁻¹ and distance to the transition state is 2.5 Å at a pulling rate of 400 nm s⁻¹ for an aqueous solution. For 30% vol vol aqueous glycerol is 1.25 × 10⁻⁹ s⁻¹ and distance to the transition state is 4.1 Å, whereas for 100% vol vol glycerol is 4.98 × 10⁻¹² s⁻¹ and distance to the transition state is 4.4 Å. The peak of the unfolding force distribution as a function of the concentration of glycerol. Mechanical stability of I27 in glycerol solutions displays a biphase force dependence for unfolding at a pulling speed of 400 nm s⁻¹ (red squares).
data confirms and extends the observations made under force-clamp conditions (Fig. 1C). Indeed, Fig. 3A shows that, whereas \( \Delta G_u \) increases linearly with the addition of glycerol, the value of \( \Delta x_u \) increases in a sharply nonlinear manner (Fig. 3B). Confidence in the values obtained through the Monte Carlo simulations is gained by comparing its estimates of \( \Delta G_u \) and \( \Delta x_u \) with those measured directly under force clamp conditions (Fig. 3A and B, circles).

Our measurements show that the size of the activation energy barrier for unfolding, \( \Delta G_u \), increases linearly with the concentration of glycerol (Fig. 3A). The increase can be as large as 11.49 kcal/mol in a pure glycerol solution. This increase is comparable to that measured for other thermally denatured proteins (10). The large increase in the activation energy barrier causes an increase in the average pulling force required to unfold the I27 protein from 200 pN (0% glycerol) up to 310 pN (100% glycerol; see 2C). The known function of protective osmolytes in vivo is to stabilize proteins against a wide variety of environmental challenges such as high pressure (42) or temperature (10). Here, we have discovered that a protective osmolyte, glycerol, greatly increases the activation energy barrier to mechanical unfolding, expanding the repertoire of known protective function of these osmolytes.

Fig. 3B now reveals the full picture of the effect of glycerol on the distance to the unfolding transition state, \( \Delta x_u \). It is clear that the interaction in \( \Delta x_u \) with glycerol concentration is highly nonlinear, and saturates at a value of \( \Delta x_u \approx 4.4 \) Å. This increase is smaller than the full 5.6 Å expected if \( \Delta x_u \) simply followed the size of the solvent molecule (see above). Nonetheless, the near doubling of \( \Delta x_u \) in pure glycerol is highly significant because it directly points to an integral structural role of solvent molecules in the unfolding transition state of the I27 protein. The unfolding transition state of a protein under force can be defined as that structure which, taken as a starting point, leads to either full unraveling or to a stable fold, with equal probability (43). Transition state structures are extremely short lived, typically requiring femtosecond laser spectroscopy for their capture (44). However, under a constant stretching force, the effect of mechanical work on the energy landscape of the unfolding protein is felt by the transition state structure, regardless of its lifetime. Thus, our finding that the distance to transition state of protein unfolding is sensitive to the size of the solvent suggests that solvent molecules are part of this short lived structure.

The steep dependency of \( \Delta x_u \) on the glycerol concentration can be understood by developing a simple model of solvent occupancy which does not require any additional information about the system, such as preferential solvation. We assume that there are \( N \) interaction sites that can be occupied by water molecules at the transition state structure, resulting in a value \( \Delta x_w \). We assume that if a single water molecule is replaced by a glycerol molecule, the transition state elongates to a value, \( \Delta x_g \). Under these conditions, the observed value of \( \Delta x_u[\text{gly}] \) for an ensemble of unfolding I27 proteins will be

\[
\Delta x_u[\text{gly}] = (P_w)^N \Delta x_w + (1 - (P_w)^N) \Delta x_g,
\]

where the probability of occupancy by a water molecule \( P_w \) is defined in terms of the glycerol concentration as \( P_w = (1 - [\text{gly}]) \), where \([\text{gly}] \) is the volume fraction of glycerol in the solution (Fig. 3). Therefore, \((P_w)^N\) is the probability that \( N \) sites are occupied by water molecules. 1 - \((P_w)^N\) corresponds to the probability that not all sites are occupied by water molecules, i.e., at least one water molecule is replaced by a glycerol molecule. If we set \( \Delta x_w = 2.5 \) Å and \( \Delta x_g = 4.4 \) Å, the measured values in pure water and pure glycerol, respectively, we can readily reproduce the steep dependency of \( \Delta x_u \) on glycerol. Fig. 3B shows plots of \( \Delta x_u[\text{gly}] \) for various values of \( N \) (\( N = 1, 3, 6, 9 \); black lines). Best fits to the measured values of \( \Delta x_u \) were obtained for \( N = 6 \) [\( x^2 = 1.5, \nu = 6, p(x^2) = 0.96 \)]. However, \( N = 7, 8 \) also had similar scores (see SI Table 1 containing \( x^2 \) and \( p(x^2) \) for all \( N \) from 1 to 10). Remarkably, the optimal values of \( N \) are similar to the known number of hydrogen bonding sites (six) between \( \beta \)-strands \( A' \) and \( G \), which are likely to be part of the unfolding transition state structure of the I27 protein (31–33, 39).

The exact nature of the transition state structure of I27 unfolding under a stretching force is unknown. Steered Molecular Dynamics (SMD) simulations can complement our AFM observations by providing a detailed atomic picture of stretching and unfolding of individual protein domains (29, 31, 32). SMD simulations involve the application of external forces to molecules in molecular dynamics simulations. The simulations are carried out by fixing one terminus of the protein, and applying external forces to the other terminus (see SI Text for details). Earlier SMD simulations showed that the simultaneous rupture of six backbone hydrogen bonds between \( \beta \)-strands \( A' \) and \( G \) of the I27 protein (Fig. 4A) was a necessary event in its mechanical unfolding (31–33). Furthermore, these simulations showed that the rupture of these interstrand hydrogen bonds could be followed by bonding to water molecules that formed bridges between the two separating strands (Fig. 4B). One way to interpret our results would be that the transition state structure
is then formed by six to eight water molecules bridging the gap between separating β-strands and taking the place of some of the broken interstrand hydrogen bonds. Our SMD simulations of forced unfolding of the I27 protein in glycerol solutions showed that the resistance to unfolding still originates from the same set of hydrogen bonds between the A’ and G β-strands. In glycerol solutions, the larger size of this cosolvent could lead to a greater gap between the separating β-strands (Fig. 4C). Given that there is a multitude of possible transition state structures formed by water, glycerol, and the protein backbone, there is no straightforward way to link a wider gap between the β-strands A’ and G in the simulations, and the experimentally measured values of Δu. In Fig. 4 B and C, we define the pulling coordinate for the separating β-strands as the distance between the first amino acid of strand A’ (Y9) and the last amino acid of strand G (K87). This distance, x(Y9)−x(K87), gets longer as the two β-strands separate under a constant force (Fig. 4D), filling the gap with solvent molecules until a transition state is reached (Fig. 4D, arrows). The elongation of the x(Y9)−x(K87) distance up to the transition state is defined as the distance to transition state, ΔXAG (Fig. 4D). The crossing of the transition state is marked by an abrupt increase in the separation length (Fig. 4D, arrows) that leads to the complete unraveling of the protein. We repeated SMD simulations of I27 unfolding measured ΔXAG in water (Fig. 4E, black bars) and in 30% glycerol solutions (Fig. 4E, red bars). The simulations show that ΔXAG increases from 2.9 ± 0.6 Å in water, up to 3.9 ± 2.1 Å in glycerol, in qualitative agreement with our observations and with the simple solvent bridging model of the unfolding transition state. From these simulations, it is not possible to determine which exact structural snapshot corresponds to the transition state of unfolding. However, it is significant that we readily find glycerol molecules between the β-strands A’ and G during the early stages of separation (Fig. 4C). It is also apparent that water molecules are found bridging β-strands A’ and G during the early stages of separation (Fig. 4C). It is interesting to consider whether a pure water transition structure might also exist in glycerol solutions. If this were true, then a glycerol molecule bridging would simply create a more stable transition state structure that is further displaced along the measured coordinate than a pure water containing transition state structure, at lower energy states. Such water molecule containing structures would then become intermediate unfolding states in glycerol solutions. However, such intermediate structures would only be ~2 Å away from the glycerol bridging transition state and thus would be difficult to capture using our current instrumentation. Additionally, the measured steep dependency of Δu on glycerol concentration (Fig. 3B) suggests that the range over which both a glycerol bridging transition state and a pure water transition state would be expected would be very narrow.

As a further test of the solvent bridging hypothesis, we replaced water with deuterium oxide (D2O), or heavy water. D2O is similar in size to water but forms hydrogen bonds that are stronger by ~0.1–0.2 kcal mol−1 at the same thermodynamic conditions of temperature and number density (45). Substituting water with D2O in our aqueous solutions increases the hydrogen bond strength of the solvent environment and enhances the stability of the protein (46). As before (Fig. 1 B and C), we measured ΔG0 and Δu in the D2O solution by fitting an Arrhenius term (19) to the force dependency of the unfolding rate (Fig. 5). Our measurements showed that, although ΔG0 increased from 23.11 kcal mol−1 in aqueous solution to 24.67 kcal mol−1 in D2O aqueous solution, Δu remained unchanged (Fig. 5B). This result lends further support to the solvent bridging model.

By combining single-molecule force spectroscopy techniques with solvent substitution, we have shown that solvent molecules form part of the mechanical transition state structure of a protein. Although we have demonstrated the crucial roles played by solvent molecules in titin, the giant elastic protein of muscle, it is likely that osmolytes also control the mechanical transition state structure of other proteins. Indeed, it will be interesting to elucidate the role of osmolytes in the mechanical transition state structure of proteins that have a distinct topology from the I27 protein. Interestingly, solvent composition is actively regulated in vivo. For example, a member of the aquaporin family of membrane channel proteins, GlpF, is highly selective for permeation of glycerol (47, 48), a naturally occurring osmolyte (26). Thus, regulation of the cellular solvent composition may be an important mechanism under conditions of mechanical stress and/or mechanical injury (34) where sustained mechanical forces applied to tissues may trigger widespread protein unfolding. A rapid compensatory increase of the cellular osmolyte concentration may therefore be “mechanically” protective.

Materials and Methods

Protein Engineering and Purification. To allow for efficient single-molecule experiments, we first constructed polyproteins using protein engineering. The details of the polyprotein engineering and purification were reported (28). Briefly, we constructed an eight-domain N-C-linked polyprotein of I27, the 27th Ig-like domain of cardiac titin, through successive cloning in modified pT7Blue vectors and then expressed the gene using vector pQE30 in Escherichia coli strain BLR(DE3).

Solvent Environment. Samples of glycerol (99%), deuterium oxide, and water were obtained from Sigma-Aldrich and used without additional purification. Solvent mixtures were prepared to obtain the desired volume fraction, volvol ratio, of the cosolvent and viscosity. Viscosities were confirmed by using a falling ball viscometer.

Single-Molecule Force Spectroscopy. We used a custom-built atomic force microscope equipped with a PicoCube P363.3-CD piezoelectric translator (Physik Instrumente, Karlsruhe, Germany) controlled by an analog PID feedback system described in refs. 16 and 28. Silicon nitride cantilevers (Veeco, Santa Barbara, CA) were calibrated for their spring constant using the equipartition theorem as reported (49). The average spring constant was ~60 pN/nm for force-extension experiments and ~15 pN/nm for force-clamp experiments. All data were obtained and analyzed using custom software written for use in Igor 5.0 (WaveMetrics, Oswego, OR). There was ~0.5 nm of peak-to-peak noise and a feedback response time of ~5 ms in all experiments. To estimate the error on our experimentally obtained rate constant, we carried out the nonparametric bootstrap method (35).

ACKNOWLEDGMENTS. We thank Arun Wiita and Sergi Garcia-Manyes for careful reading of the manuscript and Bruce Berne, Devarajan Thirumalai, and members of the Fernandez lab for helpful discussions. This work was supported by National Institutes of Health Grants HL66030 and HL61228 (to J.M.F.) and P01 AI 060915 (to H.L.).